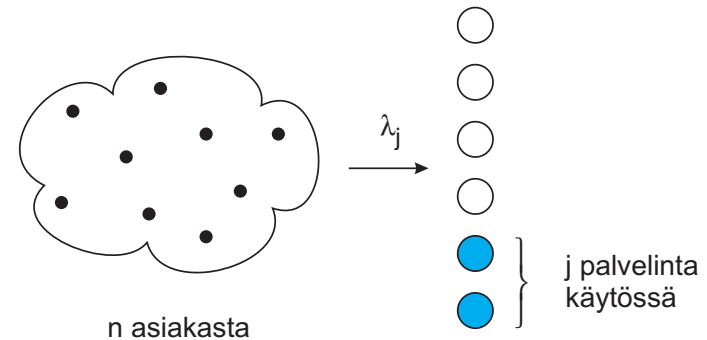


Äärellinen lähdepopulaatio: $M/M/s/s/n$ -järjestelmä

Tarkastellaan estojärjestelmää (ei odotuspaikkoja) tapauksessa, jossa saapumiset tulevat äärellisestä lähdepopulaatiosta: asiakkaiden kokonaismäärä on n .



Asiakkaan malli

- Ajatellaan konkreettisuuden vuoksi, että asiakkaat ovat esim. puhelimen käyttäjiä.
- Oletetaan, että aika seuraavaan puheluyritykseen eli asiakkaan ns. “miettimisaika” noudattaa jakaumaa $\text{Exp}(\gamma)$.
- Estynyt puheluyritys menetetään
 - ei johda uusintayritykseen
 - alkaa uusi miettimisaika: aika seuraavaan yritykseen $\sim \text{Exp}(\gamma)$
 - pitoaika $X \sim \text{Exp}(\mu)$

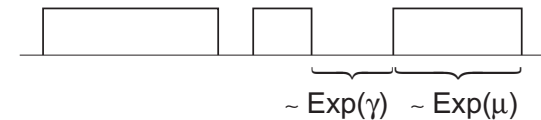
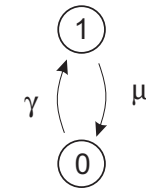
Merkitään

$$\hat{a} = \gamma/\mu$$

Yksittäisen lähteen käytäytyminen

Tapaus $s = \infty$ ($s \geq n$) Jokaiselle asiakkaalle on oma palvelin.

- Systeemissä ei ole estoa; kaikki puheluyritykset onnistuvat.
- Asiakas käyttäytyy kuvan kaksitilaisen (0,1)-mallin mukaisesti. $\text{Exp}(\mu)$ -kestoinen aktiivitila (1) ja $\text{Exp}(\gamma)$ -kestoinen miettimistila (0) vuorottelevat.



$$p_0 = \frac{1/\gamma}{1/\gamma + 1/\mu} = \frac{1}{1 + \hat{a}}, \quad p_1 = \frac{1/\mu}{1/\gamma + 1/\mu} = \frac{\hat{a}}{1 + \hat{a}}$$

- Täyden syklin (miettimisjakso + aktiivijakso) keskimääräinen kesto on $1/\gamma + 1/\mu$.
- Puheluiden saapumisnopeus (yksi per sykli) on

$$\frac{1}{1/\gamma + 1/\mu} = \frac{1/\gamma}{1/\gamma + 1/\mu} \gamma = p_0 \gamma$$

saapumisnopeus miettimistilassa γ
miettimistilan todennäköisyys p_0

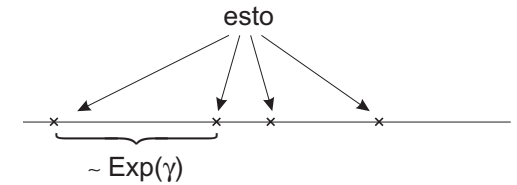
- Tarjottu kuorma = välitetty kuorma (per lähde)

$$a^\infty = p_1 = \frac{\hat{a}}{1 + \hat{a}} \quad \text{ns. tarkoitettu kuorma per lähde (ei estoa)}$$

Yksittäisen lähteen käyttäytyminen (jatkoa)

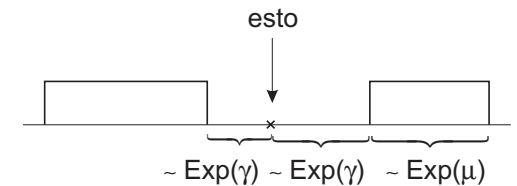
Tapaus $s = 0$: Ei yhtään palvelinta.

- Kaikki puheluyritykset estyvät.
- Puheluyrityksiä tulee frekvenssillä γ , joka on suurempi kuin tapauksessa $s = \infty$ ($p_0\gamma$), koska nyt asiakas on koko ajan miettimistilassa.



Tapaus $0 \leq s \leq n$: Varsinainen tutkittava tilanne.

- Asiakas vuorottelee $\text{Exp}(\mu)$ -kestoisen aktiivitilan ja yhdestä tai useammasta $\text{Exp}(\gamma)$ -kestoisesta jaksosta muodostuvan miettimistilan välillä.



Järjestelmän analyysi syntymä-kuolema-prosessina

Otetaan tilamuuttujaksi järjestelmässä sisällä olevien asiakkaiden (käynnissä olevien puheluiden) lukumäärä N_t .

N_t muodostaa syntymä-kuolema-tyyppisen Markovin prosessin.

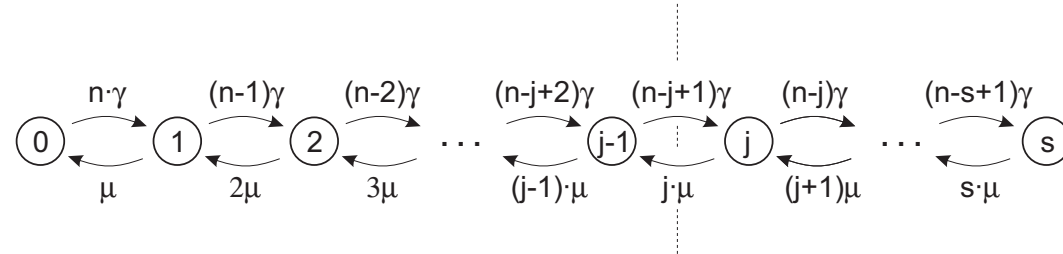
Tilamuuttuja N_t määrää (riippumatta menneisyydestä) todennäköisyydet uuden puhelun tulemiselle tai käynnissäolevan päättymiselle aikayksikköä kohden (eksponettijakaumat!).

Tilassa $N_t = j$

- $n - j$ lähdettä vapaana “miettimistilassa”
 - aika seuraavan puhelun generointiin $\sim \text{Exp}((n - j)\gamma)$
 - todennäköisyys aikayksikköä kohden siirtyä yhtä suurempaan miehitystilaan $\lambda_j = (n - j)\gamma$ (tilariippuva saapumisnopeus)
 - tilassa s kaikki palvelimet ovat käytössä, joten $\lambda_s = 0$
- j puhelua käynnissä
 - aika seuraavan päättymiseen $\sim \text{Exp}(j\mu)$
 - $\mu_j = j\mu$ (tilariippuva päättymisnopeus)

Tasapainotodennäköisyydet

Syntymä-kuolema-prosessia kuvaa seuraava tilakaavio



Merkitsemällä todennäköisyysvirrat leikkauksen yli yhtäsuuriksi, saadaan rekursio

$$(n - j + 1) \gamma \pi_{j-1} = j \mu \pi_j, \quad j = 1, 2, \dots, s$$

jonka avulla kaikki todennäköisyydet voidaan palauttaa π_0 :aan,

$$\pi_j = \frac{n - j + 1}{j} \cdot \frac{n - j + 2}{j - 1} \dots \frac{n - 1}{2} \cdot \frac{n}{1} \cdot \left(\frac{\gamma}{\mu}\right)^j \pi_0 = \binom{n}{j} \hat{a}^j \pi_0$$

Sovelletaan normiehtoa $\sum_{j=0}^s \pi_j = 1$

$$\pi_j[n] = \frac{\binom{n}{j} \hat{a}^j}{\sum_{k=0}^s \binom{n}{k} \hat{a}^k}$$

$$j = 0, 1, \dots, s$$

lähteiden lukumäärä n merkitty eksplisiittisesti näkyviin, $\pi_j[n]$

Tasapainotodennäköisyydet (jatkoa)

Merkitään $p = p_1 = a^\infty$, lähteen päälläolotodennäköisyys estottomassa tapauksessa.

$$p = \frac{\hat{a}}{1 + \hat{a}} \quad \Rightarrow \quad \hat{a} = \frac{p}{1 - p}$$

Tämän avulla tilatodennäköisyyden lauseke voidaan kirjoittaa muotoon

$$\pi_j[n] = \frac{\binom{n}{j} p^j (1 - p)^{n-j}}{\sum_{k=0}^s \binom{n}{k} p^k (1 - p)^{n-k}} \quad \begin{array}{l} j = 0, 1, \dots, s \\ (0 \text{ muulloin}) \end{array}$$

- Katkaistu binomijakauma (kun $s < n$)
- Binomijakauma, kun $s \geq n$
(estoton tapaus, kukin lähde muista riippumatta päällä todennäköisyydellä p)

Insensitiivisyys: Samalla tavalla kuin tavallisessa estojärjestelmässä tulos on insensitiivi pitoajan jakauman suhteen (vaikka nyt tehty johto nojautuikin pitoajan eksponentiaalisesti oletettuun jakaumaan).

Aika- ja kutsuestot

Aikaesto: aikaosuus, jonka systeemi viettää tilassa s ; tilan s tasapainotodennäköisyys π_s

$$E_n(s, \hat{a}) = \frac{\binom{n}{s} \hat{a}^s}{\sum_{k=0}^s \binom{n}{k} \hat{a}^k}$$

Kutsuesto: todennäköisyys, jolla saapuva kutsu estyy

- Tilariippuvasta saapumisnopeudesta johtuen tuloprosessi ei ole poissoninen.
- Kutsuesto \neq aikaesto
- Yleensäkin $\pi_j^*[n] \neq \pi_j[n]$, missä

$$\begin{cases} \pi_j^*[n] = P\{\text{kutsun saapuessa järjestelmä on tilassa } j\} \\ \pi_j[n] = P\{\text{järjestelmä on tilassa } j \text{ satunnaisella ajanhetkellä}\} \end{cases}$$

Saapuvan kutsun näkemä tilatodennäköisyys

$$\pi_j^*[n] = \frac{\lambda_j \pi_j[n]}{\sum_{k=0}^s \lambda_k \pi_k[n]}$$

Perustelu: Tarkastellaan pitkää ajanjaksoa T .

- Systemi viettää keskimäärin ajan $\pi_j[n]T$ tilassa j .
- Tänä aikana saapuu keskimäärin $\lambda_j \pi_j[n]T$ kutsua (näkevät systeemin tilassa j).
- Kaiken kaikkiaan ajassa T saapuu keskimäärin $T \sum_{k=0}^s \lambda_k \pi_k[n]$ kutsua.
- Niiden kutsujen osuus, jotka näkevät systeemin tilassa j , on lausekkeen mukainen.

Sijoittamalla ylläolevaan lausekkeeseen $\lambda_j = (n - j)\gamma$ nähdään (todistus harjoitustehtävänä)

$$\pi_j^*[n] = \pi_j[n - 1]$$

Saapuvan asiakkaan näkemä tilajakauma on sama kuin tasapainojakauma systeemissä, jossa asiakkaita on yksi vähemmän. Saapuva asiakas on ikään kuin “ulkopuolinen tarkkailija”.

Kutsuesto:

(Engsetin kaava)

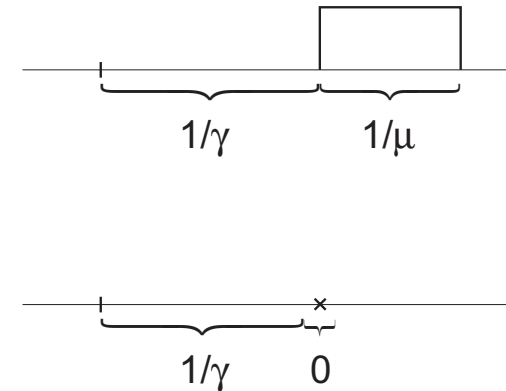
$$B_n(s, \hat{a}) = \pi_s^*[n] = \pi_s[n - 1] = \frac{\binom{n-1}{s} \hat{a}^s}{\sum_{k=0}^s \binom{n-1}{k} \hat{a}^k}$$

Toteutunut tarjottu ja välitetty kuorma

Merkitään kutsuestoa lyhyesti B :llä.

Tarkastellaan yhtä sykliä, joka sisältää

- Miettimisjakson (keskipituus $1/\gamma$)
- Palvelujakson, joka
 - todennäköisyydellä $1 - B$ on todellinen aktiivijakso (keskipituus $1/\mu$)
 - todennäköisyydellä B on surkastunut 0-kestoiseksi estotapahtumaksi



Syklin keskipituus on $1/\gamma + (1 - B) \cdot 1/\mu + B \cdot 0 = 1/\gamma + (1 - B) \cdot 1/\mu$.

Todennäköisyys, että lähde on miettimisvaiheessa on $\frac{1/\gamma}{1/\gamma + (1 - B) \cdot 1/\mu} = \frac{1}{1 + (1 - B)\hat{a}}$

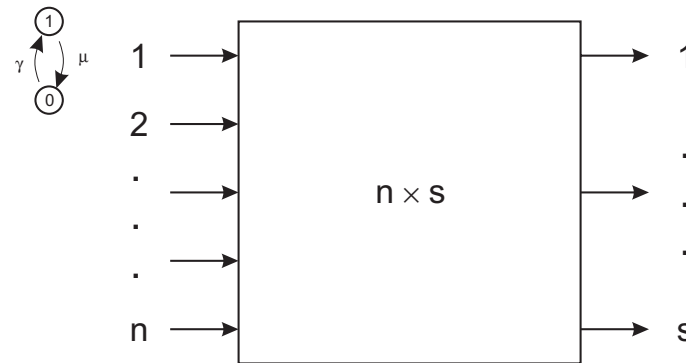
Kukin n :stä lähteestä miettimisvaiheessa ollessaan generoi uusia kutsuja nopeudella γ . Tarjottu liikenneintensiteetti a on kutsujen saapumisnopeus kertaa yhteyden pitoaika $1/\mu$.

$$a = \frac{\hat{a}}{1 + (1 - B)\hat{a}} n$$

Välitetty liikenneintensiteetti

$$a_c = (1 - B)a = \frac{(1 - B)\hat{a}}{1 + (1 - B)\hat{a}} n = \sum_{k=0}^s k \pi_k$$

Esimerkki. Äärellisen populaation järjestelmä: keskitin



- n käyttäjää, joista kukin on yhdistetty omalla johdollaan keskittimen tulopuolelle.
- $s < n$ lähtöä; liikenne on keskitetty pienempään määrään johtoja, koska on epätodennäköistä, että kaikki käyttäjät olisivat aktiivisia samanaikaisesti.
- Samanaikainen käyttö on kuitenkin mahdollista; keskityksen määrä on mitoitettava siten, että todennäköisyys estolle on riittävän pieni.
- Esto voidaan laskea annetulla kutsueston kaavalla (Engsetin kaava).